

b)

## BETWEEN-COMPOSITE INFORMATION RELEVANCE IDENTIFYING METHOD

Patent Number: JP5233719  
Publication date: 1993-09-10  
Inventor(s): OZAWA HIDEAKI; others: 01  
Applicant(s): NIPPON TELEGR & TELEPH CORP  
Requested Patent: ☐ JP5233719  
Application Number: JP19920031069 19920218  
Priority Number(s):  
IPC Classification: G06F15/40; G06F15/38  
EC Classification:  
Equivalents: JP3090233B2

---

### Abstract

---

**PURPOSE:** To attain a various relation between arbitrary multimedia information by adding a feature expressing sentence expressing each feature in a natural language to information corresponding to each media, comparing each feature quality extracted by analyzing each feature expressing sentence in the natural language, and calculating a relevance between each information.

**CONSTITUTION:** A tow link module 20 operates the relation by a keyword added to each data corresponding to the media. A retrieved result capturing module 50 receives the information retrieved by the retrieving condition output module 40, and transmits it to a content link module 60 and time link module 70. The content link module 60 operates the relation by using the feature expressing sentence of each information, and the time link module 70 operates the relation by using the generation time of each information. A control module 80 controls each module while executing an interface with a user, and prepares the final retrieved result from the information related by the content link module 60 and the time link module 70.

---

Data supplied from the esp@cenet database - I2

b)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平5-233719

(43) 公開日 平成5年(1993)9月10日

(51) Int.Cl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/40	5 2 0 D	7060-5L		
15/38	C	9194-5L		

審査請求 未請求 請求項の数4(全 9 頁)

(21) 出願番号 特願平4-31069

(22) 出願日 平成4年(1992)2月18日

(71) 出願人 000004226

日本電信電話株式会社  
東京都千代田区内幸町一丁目1番6号

(72) 発明者 小澤 英昭

東京都千代田区内幸町1丁目1番6号 日  
本電信電話株式会社内

(72) 発明者 中川 透

東京都千代田区内幸町1丁目1番6号 日  
本電信電話株式会社内

(74) 代理人 弁理士 古谷 史旺

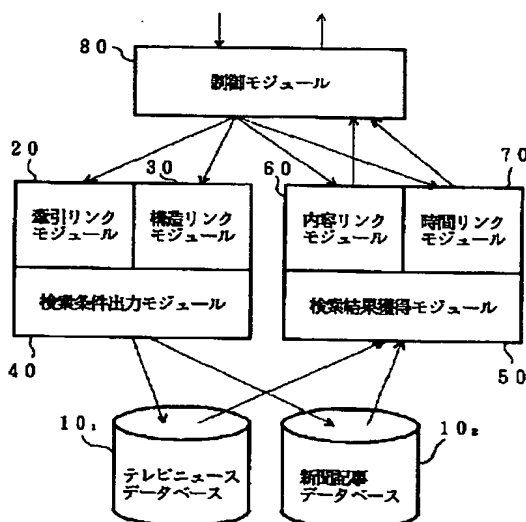
(54) 【発明の名称】 複合的な情報間の関連性識別方法

(57) 【要約】

【目的】 本発明は、映像情報、活字情報、その他の情報伝達媒体を介して提供される複合的な情報を利用するデータベースシステムにおいて、各情報間の関連性の識別に用いられる複合的な情報間の関連性識別方法に関し、任意のマルチメディア情報間に多様な関連付けを行うことができることを目的とする。

【構成】 画像情報、音声情報、テキスト情報、その他各メディア情報を含む複合的な情報間の関連性識別方法において、各情報に、個々の特徴を自然言語で表現した特徴表現文と、その情報の発生時刻データとを付加し、各情報の特徴表現文を自然言語解析して各特徴量を抽出し、各情報の発生時刻データと、各情報の特徴表現文を自然言語解析して抽出した時間に関する情報とを用いて、その情報が実際に発生した絶対時間に変換し、各情報の特徴量および絶対時間の比較により各情報間の関連度を算出することを特徴とする。

本発明方法を実現するマルチメディアデータベースシステムの実施例構成



1

## 【特許請求の範囲】

【請求項1】 画像情報、音声情報、テキスト情報、その他各メディア情報を含む複合的な情報間の関連性識別方法において、

前記各情報に、個々の特徴を自然言語で表現した特徴表現文と、その情報の発生時刻データとを付加し、

前記各情報の特徴表現文を自然言語解析して各特徴量を抽出し、

前記各情報の発生時刻データと、各情報の特徴表現文を自然言語解析して抽出した時間に関する情報とを用いて、その情報が実際に発生した絶対時間に変換し、

前記各情報の特徴量および絶対時間の比較により前記各情報間の関連度を算出することを特徴とする複合的な情報間の関連性識別方法。

【請求項2】 請求項1に記載の複合的な情報間の関連性識別方法において、

各情報の特徴表現文を自立語と修飾語の品詞ごとに分類し、品詞ごとの単語が一致する量により品詞ごとの関連度を算出し、

各情報の特徴表現文中の副詞と日時を表す語を用いて、各情報の発生時刻データを絶対時間に変換し、

前記各品詞の個数を $n$ としたときに、各品詞ごとの単語の一致量と各情報の絶対時間を $n+1$ 次元の空間上の各軸にそれぞれ割り当て、それらの大きさから前記各情報間の関連度を $n+1$ 次元の空間上の擬似的な距離として算出することを特徴とする複合的な情報間の関連性識別方法。

【請求項3】 請求項2に記載の複合的な情報間の関連性識別方法において、

各品詞ごとの単語の一致量と各情報の絶対時間を $n+1$ 次元の空間上の各軸にそれぞれ割り当てるときに、所定の軸上で扱える距離の範囲を指定して $n+1$ 次元上の距離を制御することを特徴とする複合的な情報間の関連性識別方法。

【請求項4】 請求項2に記載の複合的な情報間の関連性識別方法において、

各品詞ごとの単語が一致する量を計算する際に、特徴表現文の係受けを解析して情報の大意を表す部分とそれを修飾する部分とに分離し、各情報の修飾部分で一致する量に対して、大意を表す部分で一致する量を加重計算することを特徴とする複合的な情報間の関連性識別方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、映像として提供される情報、活字として提供される情報、その他の情報伝達媒体を介して提供される複合的な情報を利用するデータベースシステムにおいて、各情報間の関連性の識別に用いられる複合的な情報間の関連性識別方法に関する。

## 【0002】

【従来の技術】従来のマルチメディア情報を扱うデータ

2

ベースシステムでは、扱うメディアの種類や対象によってデータベース中で利用するデータの構造やデータ間の関係が異なるために、それらをデータベースの作成時に定義していた。したがって、個々のデータベースで扱える情報は、地図の情報やドキュメントの情報などのように当初に計画されていたアプリケーションに依存し、データベース固有に定義されたデータ間の関連性のもとでしか利用できなかった。

【0003】さらに、関連性の表現には主にキーワードが用いられているが、必要なデータを検索するために適当なキーワードや属性をデータベースに与える作業は、ユーザにとって大きな負担になっていた。

## 【0004】

【発明が解決しようとする課題】ところで、地図のデータベースやドキュメントのデータベースといった個々に独立に作成されたデータベース間において、例えばドキュメントのデータから地図のデータを参照したいというように、個々のメディアやデータベースを越えたマルチメディアのデータを利用したいという要求がある。

【0005】しかし、従来技術では、例えば2つのデータベース間で新たに必要となった関連性を表現するためには、新しい関係を検索することが可能なキーワードを付加した上で、新しい参照関係を示すデータベースを作成する必要があった。

【0006】本発明は、任意に作成されたマルチメディアのデータベースに対して関連性を識別する機構により、新しい参照関係が必要になっても新しい参照関係データベースを作成することなく、任意のマルチメディア情報間に多様な関連付けを行うことができる複合的な情報間の関連性識別方法を提供することを目的とする。

## 【0007】

【課題を解決するための手段】請求項1に記載の発明は、画像情報、音声情報、テキスト情報、その他各メディア情報を含む複合的な情報間の関連性識別方法において、前記各情報に、個々の特徴を自然言語で表現した特徴表現文と、その情報の発生時刻データとを付加し、前記各情報の特徴表現文を自然言語解析して各特徴量を抽出し、前記各情報の発生時刻データと、各情報の特徴表現文を自然言語解析して抽出した時間に関する情報とを用いて、その情報が実際に発生した絶対時間に変換し、前記各情報の特徴量および絶対時間の比較により前記各情報間の関連度を算出することを特徴とする。

【0008】請求項2に記載の発明は、請求項1に記載の複合的な情報間の関連性識別方法において、各情報の特徴表現文を自立語と修飾語の品詞ごとに分類し、品詞ごとの単語が一致する量により品詞ごとの関連度を算出し、各情報の特徴表現文中の副詞と日時を表す語を用いて、各情報の発生時刻データを絶対時間に変換し、前記各品詞の個数を $n$ としたときに、各品詞ごとの単語の一致量と各情報の絶対時間を $n+1$ 次元の空間上の各軸に

それぞれ割り当て、それらの大きさから前記各情報間の関連度を $n+1$ 次元の空間上の擬似的な距離として算出することを特徴とする。

【0009】請求項3に記載の発明は、請求項2に記載の複合的な情報間の関連性識別方法において、各品詞ごとの単語の一致量と各情報の絶対時間を $n+1$ 次元の空間上の各軸にそれぞれ割り当てるときに、所定の軸上で抜える距離の範囲を指定して $n+1$ 次元上の距離を制御することを特徴とする。

【0010】請求項4に記載の発明は、請求項2に記載の複合的な情報間の関連性識別方法において、各品詞ごとの単語が一致する量を計算する際に、特徴表現文の係受けを解析して情報の大意を表す部分とそれを修飾する部分とに分離し、各情報の修飾部分で一致する量に対して、大意を表す部分で一致する量を加重計算することを特徴とする。

【0011】

【作用】本発明は、各メディアに対応する情報に個々の特徴を自然言語で表現した特徴表現文を付加し、それぞれの特徴表現文を自然言語解析して抽出された各特徴量を比較することにより、各情報間の関連度を算出する。なお、自然言語解析の処理方法について、地域的な属性の抽出、人物の抽出などを変化させることにより、各情報間で複数の関連性を得ることができる。

【0012】また、各情報にその情報の発生時刻データを付加し、各情報の特徴表現文の解析と発生時刻データからその情報が実際に発生した絶対時間を割り出し、各情報を時間軸上に整列させる。

【0013】ここで、各情報の特徴量および絶対時間を比較することにより、各情報間の関連度を正確に得ることができる。すなわち、複数のデータベース間で構造やキーワードに捕らわれずに各情報の関連性を識別することができる。

【0014】

【実施例】図1は、本発明の複合的な情報間の関連性の識別方法を実現するマルチメディアデータベースシステムの実施例構成を示すブロック図である。なお、本実施例では、データベース化された新聞とテレビのニュースを例として、ある記事Aに類似した記事BやニュースCを検索する過程について説明する。

【0015】図において、本実施例のマルチメディアデータベースシステムは、それぞれテレビや新聞などの個々のメディアに対応したテレビニュースデータベース10<sub>1</sub>や新聞記事データベース10<sub>2</sub>と、牽引リンクモジュール20と、構造リンクモジュール30と、検索条件出力モジュール40と、検索結果獲得モジュール50と、内容リンクモジュール60と、時間リンクモジュール70と、制御モジュール80とにより構成される。

【0016】牽引リンクモジュール20は、各メディア対応の個々のデータに付加されたキーワードによる関連

付けを行う。構造リンクモジュール30は、各メディア対応の個々のデータの構造を操作する。検索条件出力モジュール40は、牽引リンクモジュール20および構造リンクモジュール30によって作成される検索条件で各データベース10<sub>1</sub>、10<sub>2</sub>をアクセスする。検索結果獲得モジュール50は、検索条件出力モジュール40によって検索された情報を受け、内容リンクモジュール60および時間リンクモジュール70に送出する。内容リンクモジュール60は、各情報の特徴表現文を用いて関連付けを行う。時間リンクモジュール70は、各情報の発生時刻を用いて関連付けを行う。制御モジュール80は、ユーザとのインタフェースをとりながら各モジュールを制御し、内容リンクモジュール60および時間リンクモジュール70で関連付けられた情報から最終的な検索結果を作成する。

【0017】ここで、内容リンクモジュール60および時間リンクモジュール70の構成を図2に示す。図において、内容リンクモジュール60は、検索文解析部61と、結果文解析部62と、形態素解析部63と、比較モジュール64とにより構成される。検索文解析部61は、制御モジュール80から与えられる検索対象の記事Aの特徴表現文を比較要素に基づいて解析し、その結果を比較モジュール64に与える。結果文解析部62は、検索結果獲得モジュール15を介して各データベースの検索結果を取り込み、同様に比較要素に基づいて記事B、ニュースCの特徴表現文を解析し、その結果を比較モジュール64に与える。形態素解析部63は、検索文解析部61および結果文解析部62における特徴表現文の解析を助ける。比較モジュール64は、検索文解析部61および結果文解析部62での解析結果と、制御モジュール80から与えられる閾値との比較を行い、記事B、ニュースCの各特徴量を付与する。

【0018】時間リンクモジュール70は、検索時間決定部71と、結果文解析部72と、絶対時間演算部73と、比較モジュール74とにより構成される。検索時間決定部71は、制御モジュール80から与えられる検索対象の記事Aの特徴表現文と発生時間から検索対象の基準となる時間を決め、その結果を比較モジュール74に与える。結果文解析部72は、検索結果獲得モジュール50を介して各データベースの検索結果を取り込み、記事B、ニュースCの特徴表現文と発生時刻から各記事の時間を決め、その結果を比較モジュール74に与える。絶対時間演算部73は、検索時間決定部71および結果文解析部72の各特徴表現文中から「昨日」、「今日」等といった時間情報を抽出し、その時間を年月日などによる絶対時間に変換する。比較モジュール74は、検索された記事B、ニュースCのデータが検索時間決定部71で決められた時間範囲内にあるかを判定し、記事B、ニュースCの各特徴量を付与する。

【0019】次に、制御モジュール80の構成を図3に

5

示す。図において、制御モジュール80は、検索対象文  
格納部81と、内容リンク閾値格納部82と、時間範囲  
格納部83と、解析方法格納部84と、比較要素決定部  
85と、関連度計算部86と、検索結果格納部87とに  
より構成される。検索対象文格納部81は、検索対象の  
記事を格納して内容リンクモジュール60へ与える。内  
容リンク閾値格納部82は、内容リンクモジュール60  
における特徴量計算に用いる閾値を格納して内容リンク  
モジュール60へ与える。時間範囲格納部83は、時間  
リンクモジュール70における特徴量計算に用いる時間  
範囲を格納して時間リンクモジュール70へ与える。解  
析方法格納部84は、関連度の計算アルゴリズムを格納  
して比較要素決定部85および関連度計算部86へ与え  
る。比較要素決定部85は、関連度の計算に必要な名  
詞、固有名詞などの比較要素を決定して内容リンクモ  
ジュール60へ与える。関連度計算部86は、内容リンク  
モジュール60および時間リンクモジュール70によっ  
て決定された特徴量を取り込み、関連度の計算アルゴ  
リズムを用いて関連度を算出して検索結果格納部87に与  
える。検索結果格納部87は、最終的に関連度ありと判  
断されたデータを格納する。

【0020】以下、記事Aに類似した記事Bやニュース  
Cを検索する過程の動作手順について、以上示した各部  
の構成と、図4に示すフローチャートを参照して説明す  
る。まず、ユーザは制御モジュール80に、検索対象の  
記事Aと、関連度計算の解析方法のアルゴリズムと、時  
間範囲と、閾値とを設定することにより、制御モジュ  
ール80は動作を開始する。

【0021】制御モジュール80は、比較要素決定部8  
5で関連度の計算に必要な比較要素を決定し、記事Aの  
特徴表現文、閾値とともに内容リンクモジュール60に  
設定する(ステップ1)。

【0022】次に、制御モジュール80は、記事Aの特  
特徴表現文、発生時間および時間範囲を時間リンクモ  
ジュール70に設定する(ステップ2)。次に、制御モ  
ジュール80は、各データベースをキーワード検索するた  
めのキーワードを牽引リンクモジュール20に設定する  
(ステップ3)。さらに、例えば新聞記事から写真の部  
分のみを得るといった構造を操作するための条件、すな  
わち抽出したい構造表現を構造リンクモジュール30に  
設定する(ステップ4)。

【0023】検索条件出力モジュール40では、牽引リ  
ンクモジュール20および構造リンクモジュール30に  
設定された条件を用いて、テレビニュースデータベース  
10、および新聞記事データベース10、を検索する  
(ステップ5)。検索された結果は、検索結果獲得モ  
ジュール50を介して内容リンクモジュール60および時  
間リンクモジュール70に送られる。なお、検索結果獲  
得モジュール50は、ここではバッファの役割を果たし  
ている。

6

【0024】内容リンクモジュール60では、検索文解  
析部61で記事Aの特徴表現文について形態素解析した  
後に、比較要素に従って解析に必要なデータを生成し、  
比較モジュール64に格納する。また、結果文解析部6  
2では、検索された記事B、ニュースCの特徴表現文に  
ついて形態素解析した後に、比較要素に従って比較対  
象のデータを生成し、比較モジュール64に格納する。比  
較モジュール64では、記事Aと記事B、ニュースCの  
各解析データを比較し、閾値よりも多くの一致があるデ  
ータには0.00以上1.00未満の特徴量を付加し、閾値よ  
りも一致が少ないデータには1.00の特徴量を付加し、各記  
事のデータとともに制御モジュール80に送出する(ス  
テップ6)。

【0025】時間リンクモジュール70では、検索時間  
決定部71で記事Aの特徴表現文および発生時間を用い  
て、例えば「昨日」という時間情報を「1991年9月10  
日」のような絶対時間に変換し、時間範囲によって「19  
91年8月から1991年10月」のような時間の区間に変換  
し、比較モジュール74に格納する。また、結果文解析  
部72では、検索された記事B、ニュースCの特徴表現  
文から、例えば事件のあった日付、報道された日付のデ  
ータを抽出して比較モジュール74に格納する。比較モ  
ジュール74では、記事Aと記事B、ニュースCの各時  
間データを比較し、時間範囲内に存在するデータには0.  
00以上1.00未満の特徴量を付加し、時間範囲外のデー  
タには1.00の特徴量を付加し、制御モジュール80に送出  
する(ステップ7)。

【0026】制御モジュール80では、関連度計算部8  
6に内容リンクモジュール60および時間リンクモジュ  
ール70によって決定された特徴量を取り込み、与えら  
れた解析方法を用いて関連度を算出する(ステップ  
8)。関連性判定の一例としては、内容リンクモジュ  
ール60から得られた特徴量を $\alpha_i$  ( $i=1\sim n$ )、時間リ  
ンクモジュール70から得られた特徴量を $\beta$ とし、それ  
らを擬似的な距離と見なして、

【0027】

【数1】

$$\sqrt{\sum_{i=1}^n (\alpha_i)^2 + \beta^2} \quad \dots (1)$$

の計算式により、相対的な距離を計算する。ここで、解  
析方法中に指示された所定の距離内にあるデータに関し  
てのみ関連性ありと判断する。

【0028】検索結果格納部87では、関連性ありと判  
断されたデータについて、相対的な距離から関連度の高  
い順に並べ替えて格納する(ステップ9)。ここで、特  
特徴表現文、データ構造、関連度計算のアルゴリズムにお  
ける空間、時間リンクモジュール70における特徴量の  
決め方、内容リンクモジュール60で得られる特徴量の  
決め方についてその一例を説明する。

7

【0029】特徴表現文は、新聞や雑誌のような活字のメディアでは本文のデータ、テレビやラジオのような映像・音声のメディアではスクリプトを用いることにより実現できる。

【0030】データ構造は、テレビニュースを例にとると、映像、音声、特徴表現文、放映日時およびキーワードからなる。キーワードには、例えば日本などの地名や経済などのジャンルといったものが付けられる。

【0031】関連度計算のアルゴリズムにおける空間は、内容リンクモジュール60で生成される特徴量の中で、特徴表現文の中に多数存在し情報を大まかに示している名詞による特徴量と、動作の状態を示すサ変名詞による特徴量と、事件のあった場所や関わった人物などを示す固有名詞による特徴量と、「昨日」といった時間を表す副詞や「4月」といった時間を表す数詞から生成される特徴量とを用い、4次元の空間として表現する。 \*

$$\sqrt{\sum_{i=1}^n (\mu_i \cdot \alpha_i)^2 + (\mu_B \cdot \beta)^2} \quad \dots(2)$$

とする。ただし、 $\mu_i$ 、 $\mu_B$  は、0.00以上の実数で、標準は1.0である。この場合、例えば固有名詞の軸強度が0.010であったとすると、情報Xとの特徴量について、名詞の特徴量が0.30、サ変名詞の特徴量が0.60、固有名詞の特徴量が0.20であり、時間リンクモジュール70で得られた特徴量が0.00である情報Yの擬似的な距離は0.67となる。また、名詞の特徴量が0.30、サ変名詞の特徴量が0.60、固有名詞の特徴量が0.60であり、時間リンクモジュール70で得られた特徴量が0.00である情報Zの擬似的な距離も0.67となる。したがって、情報Yおよび情報Zの情報Xに対する類似度は同じと判定される。このように、軸強度を変えることで、例えば固有名詞の軸の影響を弱めれば、サ変名詞が表現する事件が類似しているといった観点での類似度が強調され、逆に固有名詞の軸の影響を強めれば人物や場所などの類似が強調され、関連性の識別に多様性を与えることができる。

【0035】時間リンクモジュール70における特徴量の決め方は、時間リンクモジュール70が記事Aと同日のみの情報しか扱わない場合を0.00とし、無限に古い情報までを扱う場合を1.00として、時間幅を双曲線関数的に表現することで達成できる。たとえば、記事Aとの時間差Xを日数で表し、特徴量Yを

$$Y=1.00-1/(X+1) \quad \dots(3)$$

と表すと、記事Aと同日の記事について時間リンクモジュール70で抽出される特徴量は0.00となり、1年前の情報は0.997となる。

【0036】特徴量Yの他の決め方は、メディアの種類に応じて時間リンクモジュール70が特徴量を決定すれば、テレビと新聞とのように異なるメディアに対しては、情報が発生してから報道されるまでの時間差を吸収

8

\*【0032】例えば、情報Xとの特徴量について、名詞の特徴量が0.30、サ変名詞の特徴量が0.60、固有名詞の特徴量が0.20であり、時間リンクモジュール70で得られた特徴量が0.00である情報Yに対して、動作手順のステップ8で述べた(1)式により、擬似的な距離は0.70となる。また、名詞の特徴量が0.30、サ変名詞の特徴量が0.60、固有名詞の特徴量が0.60であり、時間リンクモジュール70で得られた特徴量が0.00である情報Zに対して、擬似的な距離は0.90となる。したがって、情報Yの方が情報Xに類似した情報と判定される。

【0033】この他に各軸の特徴量を変化させることにより、擬似的な距離を変化させることも可能である。たとえば、(1)式に各軸の影響度を示す軸強度 $\mu_i$  ( $i=1 \sim n$ )、 $\mu_B$ を加えて、

【0034】

【数2】

0に入力された記事Aが午後のテレビニュースであった場合に、新聞で翌日の朝刊の記事に対しては時間リンクモジュール70で得られる特徴量を0.00とし、他の情報に対しては双曲線関数で表現する。また、時間リンクモジュール70に入力された記事Aが午前のテレビニュースであれば新聞の夕刊の記事、記事Aが新聞の朝刊であれば前日のテレビニュース、記事Aが新聞の夕刊であればお昼までのテレビニュースに対しては、時間リンクモジュール70で得られる特徴量を0.00とする。この結果、新聞とテレビニュースとの間にあるメディアによる時間差を吸収することができ、ある記事に最も類似した記事を検索し易くすることができる。

【0037】内容リンクモジュール60で得られる特徴量の決め方は、検索対象の記事Aの特徴表現文から得られた比較対象データと、データベースからの検索文から得られた単語とが完全に一致する場合を0.00とし、まったく一致しない場合を1.00としたときに、一致した単語数をa、比較対象データの単語数をbとすると、その特徴量Wは、

$$W=1.00-a/b \quad \dots(4)$$

として表現される。たとえば、記事Aにおいて名詞が10個存在したときに、検索した記事中に一致する名詞が8個存在すれば、特徴量は0.20となる。なお、内容リンクモジュール60の閾値は外部から設定される。

【0038】特徴量Wの他の決め方は、内容リンクモジュール60の特徴量の計算において特徴表現文を係受け解析し、文の主題を表現する文節と、それを修飾する文節に分離する。すなわち、内容リンクモジュール60に入力された記事Aと、データベースから得られた記事のそれぞれの特徴表現文を係受け解析し、主題(大意)を表す文節から得られた単語と修飾する文節から得られた

単語とを分類する。例えば比較対象データにおいて、主題を表す文節から名詞が10個得られ、検索した記事中の主題を表す文節中に一致する名詞が7個あり、修飾する文節どうしでは6個中3個の名詞が一致したとすると、修飾する文節の影響を0.30、主題の文節の影響を0.70とすれば、特徴量Wは、 $0.70 \times 0.70 + 0.30 \times 0.50 = 0.36$ となる。この結果、副次的情報による関連性の計算への悪影響を排除することができる。

【0039】

【発明の効果】以上説明したように、本発明の複合的な情報間の関連性識別方法では、例えば新聞記事とテレビニュースのような異なるメディアの情報間に存在する関連性を容易に識別することができる。すなわち、それぞれ異なる作成者により構築されたマルチメディアデータベースに対して、高機能の情報検索サービスを提供することができる。

【0040】さらに、本発明方法では、個々の情報に特徴表現文と発生時刻データとを付加するだけでよく、複数のデータベースに渡ってキーワードなどの関連性を考慮する必要がなくなり、データベース作成者への負担を大幅に軽減することができる。

【図面の簡単な説明】

【図1】本発明の複合的な情報間の関連性の識別方法を実現するマルチメディアデータベースシステムの実施例構成を示すブロック図である。

【図2】内容リンクモジュール60および時間リンクモジュール70の構成例を示すブロック図である。

【図3】制御モジュール80の構成例を示すブロック図

である。

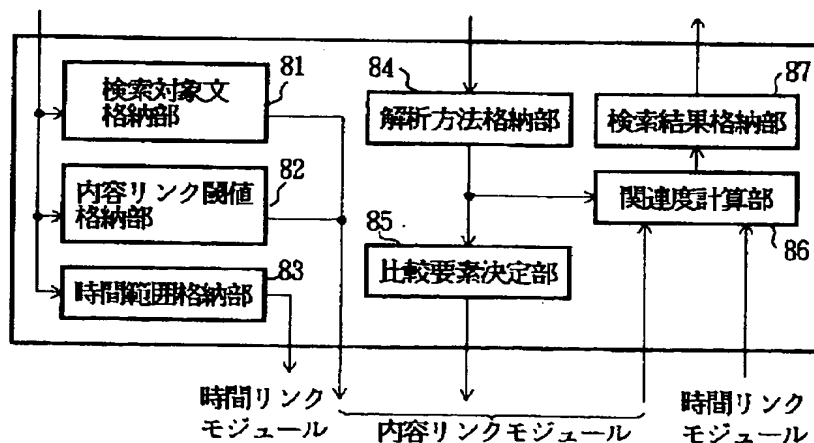
【図4】本発明実施例の動作手順を説明するフローチャートである。

【符号の説明】

- 10: テレビニュースデータベース
- 10: 新聞記事データベース
- 20 牽引リンクモジュール
- 30 構造リンクモジュール
- 40 検索条件出力モジュール
- 50 検索結果獲得モジュール
- 60 内容リンクモジュール
- 61 検索文解析部
- 62 結果文解析部
- 63 形態素解析部
- 64 比較モジュール
- 70 時間リンクモジュール
- 71 検索時間決定部
- 72 結果文解析部
- 73 絶対時間演算部
- 74 比較モジュール
- 80 制御モジュール
- 81 検索対象文格納部
- 82 内容リンク閾値格納部
- 83 時間範囲格納部
- 84 解析方法格納部
- 85 比較要素決定部
- 86 関連度計算部
- 87 検索結果格納部

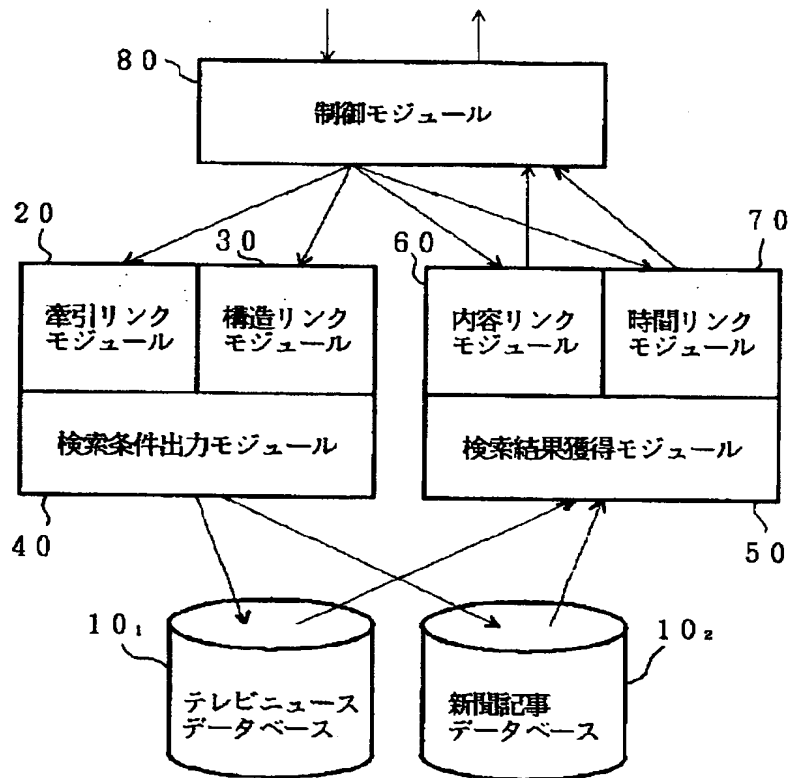
【図3】

制御モジュール80の構成例



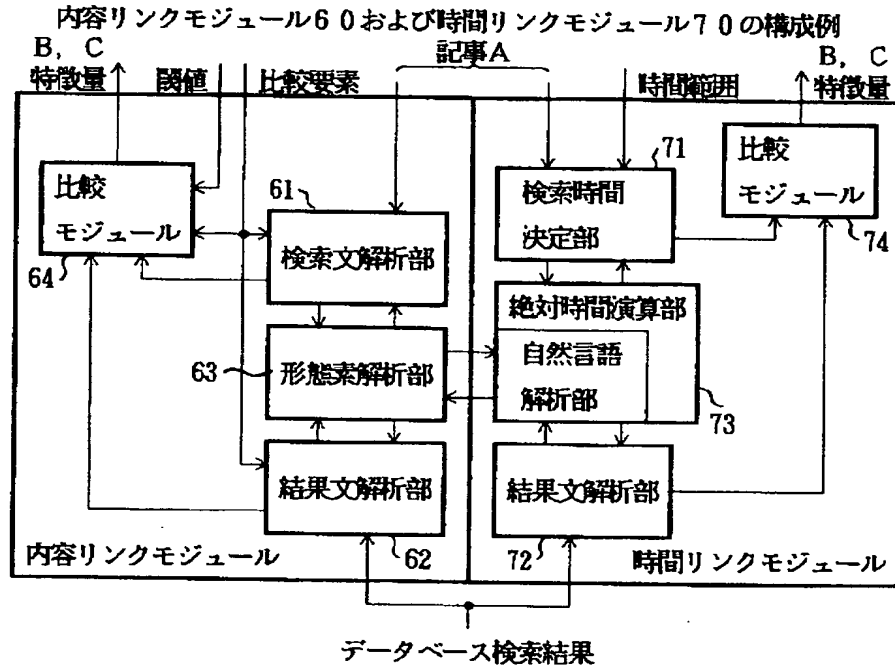
【図1】

本発明方法を実現するマルチメディアデータベースシステムの実施例構成





【図2】



【図4】

## 本発明実施例の動作手順

